

**STIC Biotechnology Systems Branch**

**RAW SEQUENCE LISTING**  
**ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/532,681

Source: PCT

Date Processed by STIC: 1-23-06

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/efc/efs/downloads/documents.htm>) , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

## Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	SERIAL NUMBER: <u>10/532,681</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input checked="" type="checkbox"/> Wrapped Nucleics Wrapped Aminos	The <u>number/text</u> at the <u>end</u> of each line "wrapped" down to the next line. <u>This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."</u>	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line <b>not exceed</b> 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do <b>not</b> use tab codes between numbers; use <b>space characters</b> , instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was <b>not</b> saved in ASCII(DOS) text, as <b>required</b> by the Sequence Rules. <b>Please ensure your subsequent submission is saved in ASCII text.</b>	
5 <input type="checkbox"/> Variable Length	Sequence(s) <u>      </u> contain n's or Xaa's representing more than one residue. <b>Per Sequence Rules, each n or Xaa can only represent a single residue.</b> Please present the <b>maximum</b> number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) <u>      </u> . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) <u>      </u> missing. If intentional, please insert the following lines for <b>each</b> skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to <b>include</b> the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) <u>      </u> missing. If <b>intentional</b> , please insert the following lines for <b>each</b> skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is <b>MANDATORY</b> if n's or Xaa's are present. In <220> to <223> section, please explain location of <b>n</b> or <b>Xaa</b> , and which residue <b>n</b> or <b>Xaa</b> represents.	
10 <input type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only <b>valid</b> <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is <b>required</b> when <213> response is Unknown or is Artificial Sequence	
11 <input checked="" type="checkbox"/> Use of <220>	Use of <u>&lt;220&gt; to &lt;223&gt;</u> is <b>MANDATORY</b> if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <u>&lt;220&gt; to &lt;223&gt;</u> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n/Xaa	<b>"n"</b> can <b>only</b> represent a single <u>nucleotide</u> ; <b>"Xaa"</b> can <b>only</b> represent a single <u>amino acid</u>	



PCT

## RAW SEQUENCE LISTING

DATE: 01/23/2006

PATENT APPLICATION: US/10/532,681

TIME: 09:12:04

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

3 <110> APPLICANT: LUKYANOV, Sergei Anatolievich  
 4 SHAGIN, Dmitry Alexeevich  
 5 YANUSHEVICH, Yury Grigorievich  
 7 <120> TITLE OF INVENTION: FLUORESCENT PROTEINS AND CHROMOPROTEINS FROM NON-AEQUOREA HYDROZOA  
 8 SPECIES AND METHODS FOR USING SAME  
 10 <130> FILE REFERENCE: U 015745-9  
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/532,681  
 11 <141> CURRENT FILING DATE: 2005-04-26  
 13 <160> NUMBER OF SEQ ID NOS: 22  
 15 <170> SOFTWARE: PatentIn version 3.1

## ERRORED SEQUENCES

Does Not Comply  
 Corrected Diskette Needed  
 (pg. 1-5) ←

134 <210> SEQ ID NO: 5  
 135 <211> LENGTH: 705  
 136 <212> TYPE: DNA  
 137 <213> ORGANISM: Artificial sequence  
 W--> 138 <220> FEATURE:  
 139 <223> OTHER INFORMATION: phiYFP-M0 mutant of the phiYFP  
 W--> 140 <400> SEQUENCE: 5  
 142 atgcctagtg gagcactggt gttccacgga aagatcccat atgttggtga gatggagggg 60  
 143 aatgttgatg gacacacatt ctccattaga ggtaaagggt atggagatgc aagtgttggt 120  
 144 aaagttgatg cccaattcat ctgcacaact ggagatgtac cagttccatg gtcaacttta 180  
 145 gtaacaacac ttaacttatgg tgaacaatgc ttgcaccaat atggtccaga attaaaggat 240  
 E--> 146  
 ttctacaaga gttgcatgcc tgaaggctat gtgcaggagc gtacaatcac atttgaaggg 300gacggaaact t  
 147 gtcaaaactta atggacaagg atttaagaaa gacggacatg tgcttggaag gaatcttgaa 420  
 148 ttcaatttca cacctcattg tctttacatt tggggagatc aggcataatca tggtttgaag 480  
 149 tctgttttca aaattcgcca tgagattact ggatcaag gagacttcat tgttcagac 540  
 150 cacacccaaa tgaacacacc cattggtggt ggaccagtcc atgtccctga aaaccatcat 600  
 151 atgagctacc atgtcaagct cagcaaagat gttactgatc acagggataa catgagcttg 660  
 152 aaggaaaccg tacgggctgt ggattgcaga aaaacatatc tttaa 705  
 275 <210> SEQ ID NO: 10  
 276 <211> LENGTH: 234  
 277 <212> TYPE: PRT  
 278 <213> ORGANISM: Artificial sequence  
 W--> 279 <220> FEATURE:  
 280 <223> OTHER INFORMATION: humanized version of the phiYFP-M1  
 W--> 281 <400> SEQUENCE: 10  
 283 Met Ser Ser Gly Ala Leu Leu Phe His Gly Lys Ile Pro Tyr Val Val  
 284 1 5 10 15  
 285 Glu Met Glu Gly Asn Val Asp Gly His Thr Phe Ser Ile Arg Gly Lys

See item # 1 on error  
 Summary sheet.

286

20

25

30

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,681

DATE: 01/23/2006

TIME: 09:12:04

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

```

287 Gly Tyr Gly Asp Ala Ser Val Gly Lys Val Asp Ala Gln Phe Ile Cys
288           35                      40                      45
289 Thr Thr Gly Asp Val Pro Val Pro Trp Ser Thr Leu Val Thr Thr Leu
290           50                      55                      60
292 Thr Tyr Gly Ala Gln Cys Phe Ala Lys Tyr Gly Pro Glu Leu Lys Asp
293 65                      70                      75                      80
294 Phe Tyr Lys Ser Cys Met Pro Asp Gly Tyr Val Gln Glu Arg Thr Ile
295                      85                      90                      95
296 Thr Phe Glu Gly Asp Gly Asn Phe Lys Thr Arg Ala Glu Val Thr Phe
297                      100                     105                     110
298 Glu Asn Gly Ser Val Tyr Asn Arg Val Lys Leu Asn Gly Gln Gly Phe
299           115                      120                      125
300
Lys Lys Asp Gly His Val Leu Gly Lys Asn Leu Glu Phe Asn Phe Thr      130
301 Pro His Cys Leu Tyr Ile Trp Gly Asp Gln Ala Asn His Gly Leu Lys
E--> 302 145                      150                      155                      160
303 Ser Ala Phe Lys Ile Cys His Glu Ile Thr Gly Ser Lys Gly Asp Phe
E--> 304                      165                      170                      175
305 Ile Val Ala Asp His Thr Gln Met Asn Thr Pro Ile Gly Gly Gly Pro
E--> 306                      180                      185                      190
307 Val His Val Pro Glu Tyr His His Met Ser Tyr His Val Lys Leu Ser
E--> 308                      195                      200                      205
309 Lys Asp Val Thr Asp His Arg Asp Asn Met Ser Leu Lys Glu Thr Val
E--> 310                      210                      215                      220
311 Arg Ala Val Asp Cys Arg Lys Thr Tyr Leu
E--> 312 225                      230
406 <210> SEQ ID NO: 14
407 <211> LENGTH: 232
408 <212> TYPE: PRT
409 <213> ORGANISM: hydromedusa 2 from sub-order Anthomedusae
W--> 410 <400> SEQUENCE: 14
412 Met Glu Gly Gly Pro Ala Leu Phe Gln Ser Asp Met Thr Phe Lys Ile
413 1                      5                      10                      15
414 Phe Ile Asp Gly Val Val Asn Asp Gln Lys Phe Thr Ile Ile Ala Asp
415           20                      25                      30
416 Gly Ser Ser Lys Phe Pro His Gly Asp Phe Asn Val His Ala Val Cys
417           35                      40                      45
418 Glu Thr Gly Lys Leu Pro Met Ser Trp Lys Pro Ile Cys His Leu Ile
419           50                      55                      60
420 Gln Tyr Gly Glu Pro Phe Phe Ala Lys Tyr Pro Asn Gly Ile Ser His
421 65                      70                      75                      80
422 Phe Ala Gln Glu Cys Phe Pro Glu Gly Leu Thr Ile Asp Arg Thr Val
423           85                      90                      95
424 Arg Phe Glu Asn Asp Gly Thr Met Thr Ser His His Thr Tyr Glu Leu
425           100                     105                     110
426 Asp Gly Thr Cys Val Ile Ser Arg Ile Thr Val Asn Cys Asp Gly Phe
427           115                      120                      125
428 Gln Pro Asp Gly Pro Ile Met Lys Asp Gln Leu Val Asp Ile Leu Pro
429           130                      135                      140
430 Thr Glu Thr His Met Phe Pro His Gly Ser Asn Ala Val Arg Gln Leu

```

See item  
#1 on error  
summary  
sheet.

Insert on line  
62207-62237



## RAW SEQUENCE LISTING

DATE: 01/23/2006

PATENT APPLICATION: US/10/532,681

TIME: 09:12:04

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

647 &lt;212&gt; TYPE: PRT

648 &lt;213&gt; ORGANISM: Artificial sequence

W--&gt; 649 &lt;220&gt; FEATURE:

650 &lt;223&gt; OTHER INFORMATION: humanized S3-2 mutant

W--&gt; 651 &lt;400&gt; SEQUENCE: 22

653 Met Glu Gly Gly Pro Ala Leu Phe Gln Ser Asp Met Thr Phe Lys Ile  
654 1 5 10 15  
656 Phe Ile Asp Gly Val Val Asn Gly Gln Lys Phe Thr Ile Val Ala Asp  
657 20 25 30  
658 Gly Ser Ser Lys Phe Pro His Gly Asp Phe Asn Val His Ala Val Cys  
659 35 40 45  
660 Glu Thr Gly Lys Leu Pro Met Ser Trp Lys Pro Ile Cys His Leu Ile  
661 50 55 60  
662 Gln Tyr Gly Glu Pro Phe Ala Arg Tyr Pro Asn Gly Ile Ser His  
663 65 70 75 80  
664 Phe Ala Gln Glu Cys Phe Pro Glu Gly Leu Ser Ile Asp Arg Thr Val  
665 85 90 95  
666 Arg Phe Glu Asn Asp Gly Thr Met Thr Ser His His Thr Tyr Glu Leu  
667 100 105 110  
668 Asp Gly Thr Cys Val Val Ser Arg Ile Thr Val Asn Cys Asp Gly Phe  
669 115 120 125  
670 Gln Pro Asp Gly Pro Ile Met Arg Asp Gln Leu Val Asp Ile Leu Pro  
671 130 135 140  
672 Asn Glu Thr His Met Phe Pro His Gly Pro Asn Ala Val Arg Gln Leu  
673 145 150 155 160  
674 Ala Phe Ile Gly Phe Thr Thr Ala Asp Gly Gly Leu Met Met Ser His  
675 165 170 175  
676 Phe Asp Ser Lys Met Thr Phe Asn Gly Ser Arg Ala Ile Lys Ile Pro  
677 180 185 190  
678 Gly Pro His Phe Val Thr Thr Ile Thr Lys Gln Met Lys Asp Thr Ser  
679 195 200 205  
680 Asp Lys Arg Asp His Val Cys Gln Arg Glu Val Thr Tyr Ala His Ser  
681 210 215 220  
682 Val Pro Arg Ile Thr Ser Ala Ile  
683 225 230

E--&gt; 687 1

*pls delete*

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 01/23/2006

PATENT APPLICATION: US/10/532,681

TIME: 09:12:05

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:5; Line(s) 146

Seq#:10; Line(s) 300

Seq#:14; Line(s) 431

Seq#:15; Line(s) 456

Seq#:19; Line(s) 577



## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/532,681

DATE: 01/23/2006

TIME: 09:12:05

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

L:11 M:270 C: Current Application Number differs, Missing <140> CURRENT APPLICATION NUMBER: is Added.

L:16 M:283 W: Missing Blank Line separator, <210> field identifier

L:20 M:283 W: Missing Blank Line separator, <400> field identifier

L:40 M:283 W: Missing Blank Line separator, <400> field identifier

L:78 M:283 W: Missing Blank Line separator, <220> field identifier

L:80 M:283 W: Missing Blank Line separator, <400> field identifier

L:99 M:283 W: Missing Blank Line separator, <220> field identifier

L:101 M:283 W: Missing Blank Line separator, <400> field identifier

L:138 M:283 W: Missing Blank Line separator, <220> field identifier

L:140 M:283 W: Missing Blank Line separator, <400> field identifier

L:146 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:3 ✓

L:158 M:283 W: Missing Blank Line separator, <220> field identifier

L:160 M:283 W: Missing Blank Line separator, <400> field identifier

L:198 M:283 W: Missing Blank Line separator, <220> field identifier

L:200 M:283 W: Missing Blank Line separator, <400> field identifier

L:219 M:283 W: Missing Blank Line separator, <220> field identifier

L:221 M:283 W: Missing Blank Line separator, <400> field identifier

L:258 M:283 W: Missing Blank Line separator, <220> field identifier

L:260 M:283 W: Missing Blank Line separator, <400> field identifier

L:279 M:283 W: Missing Blank Line separator, <220> field identifier

L:281 M:283 W: Missing Blank Line separator, <400> field identifier

L:302 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:304 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:306 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:308 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:310 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:312 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10

L:312 M:252 E: No. of Seq. differs, <211> LENGTH:Input:234 Found:218 SEQ:10 ✓ OK

L:318 M:283 W: Missing Blank Line separator, <400> field identifier

L:343 M:283 W: Missing Blank Line separator, <400> field identifier

L:384 M:283 W: Missing Blank Line separator, <400> field identifier

L:410 M:283 W: Missing Blank Line separator, <400> field identifier

L:431 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:432 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:434 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:436 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:438 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:440 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14

L:446 M:283 W: Missing Blank Line separator, <220> field identifier

L:448 M:283 W: Missing Blank Line separator, <400> field identifier

L:456 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:3

L:466 M:283 W: Missing Blank Line separator, <220> field identifier

L:468 M:283 W: Missing Blank Line separator, <400> field identifier

L:505 M:283 W: Missing Blank Line separator, <220> field identifier

L:508 M:283 W: Missing Blank Line separator, <400> field identifier

L:526 M:283 W: Missing Blank Line separator, <220> field identifier

L:529 M:283 W: Missing Blank Line separator, <400> field identifier

L:567 M:283 W: Missing Blank Line separator, <220> field identifier

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/532,681

DATE: 01/23/2006

TIME: 09:12:05

Input Set : A:\u0157459 sequence listing.txt

Output Set: N:\CRF4\01232006\J532681.raw

L:570 M:283 W: Missing Blank Line separator, <400> field identifier  
L:577 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:3  
L:588 M:283 W: Missing Blank Line separator, <220> field identifier  
L:591 M:283 W: Missing Blank Line separator, <400> field identifier  
L:628 M:283 W: Missing Blank Line separator, <220> field identifier  
L:630 M:283 W: Missing Blank Line separator, <400> field identifier  
L:649 M:283 W: Missing Blank Line separator, <220> field identifier  
L:651 M:283 W: Missing Blank Line separator, <400> field identifier  
L:687 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:22